

## APPENDIX

### MARKED UP CLAIMS PURSUANT TO 37 CFR 1.121(c)(ii) TO ACCOMPANY THE RESPONSE AND AMENDMENT

1. (Amended) A system for measuring an ability of a subject, comprising:
  - a first set of task items that require the subject to provide one or more spoken responses;
  - a speech recognition system coupled to receive the spoken response and to provide an estimate of the spoken response, the speech recognition system having an associated accuracy;
  - a scoring device, the scoring device being operable to convert the estimate into an item score; and
  - a computation device, the computation device providing a subject score based on a combination of item scores using a scoring computation model that depends upon an expected item-dependent operating characteristic of the speech recognition system, including the associated accuracy of the speech recognition system.
  
8. (Amended) A method for measuring an ability of a subject, comprising:
  - providing a set of task items;
  - generating a difficulty value for each task item in the set, the difficulty value being based upon the task item and a performance measurement associated with an automatic device that measures task performance, wherein the performance measurement comprises a measure of an ability of the automatic device to accurately recognize responses to the set of tasks;
  - obtaining a response to each task item from the subject; and
  - combining the difficulty values and the responses from the subject to form a subject score reflecting at least one of a linguistic ability and one cognitive ability of the subject.

14. (Amended) A method for measuring an ability of a subject comprising:  
providing a set of tasks and a device that automatically measures performance of the tasks;  
determining a difficulty value for each task, wherein the difficulty value is based upon the task and upon a performance measure associated with an ability of the automated device to accurately assess performance of the task;  
obtaining verbal responses to the tasks from the subject; and  
combining the verbal responses and the difficulty values to form a subject score reflecting at least one of a linguistic ability and a cognitive ability of the subject.

16. (Amended) An apparatus for determining a difficulty value of items in a test, comprising:  
a set of responses to the items from a number of individuals;  
an automated grader, wherein the automatic grades receives the set of responses and provides graded responses; and  
means for reducing the graded responses to a set of item difficulties, said item difficulties normalizing the items by reflecting an ability of the automatic grader to accurately grade the set of responses.

17. (Amended) A method for determining a difficulty value of items in a text, comprising:

obtaining a set of responses to the items from a number of individuals;  
automatically grading the set of responses, thereby generating graded responses; and  
reducing the graded responses to a set of item difficulties, said item difficulties including  
a measurement of accuracy for the act of automatically grading the set of responses for purposes  
of normalizing the items to provide an accurate assessment.